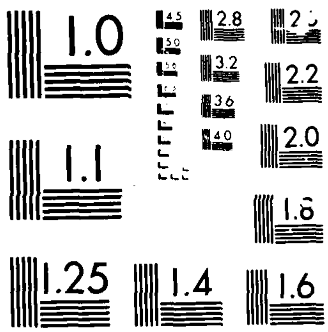
AD-A169 606 CONDITIONAL BOUNDARY CROSSING PROBABILITIES WITH APPLICATIONS TO CHANGE-POINT PROBLEMS(U) STANFORD UNIV CA DEPT OF STATISTICS B JAMES ET AL. JUN 86 TR-36 1/1

UNCLASSIFIED N00014-77-C-0306 F/G 12/1 NL


1.0
4.5
5.0
5.6
2.8
3.2
3.6
4.0
2.2
2.0
1.1
1.8
1.25
1.4
1.6

CONDITIONAL BOUNDARY CROSSING PROBABILITIES WITH APPLICATIONS TO CHANGE-POINT PROBLEMS

by

Barry James
and
Kang Ling James
IMPA, Rio de Janeiro

David Siegmund
Stanford University

TECHNICAL REPORT NO. 36
JUNE 1986

PREPARED UNDER CONTRACT
N00014-77-C-0306 (NR-042-373)
FOR THE OFFICE OF NAVAL RESEARCH

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

CONDITIONAL BOUNDARY CROSSING PROBABILITIES WITH APPLICATIONS TO CHANGE-POINT PROBLEMS

by

Barry James

and

Kang Ling James

IMPA, Rio de Janeiro

David Siegmund

Stanford University

TECHNICAL REPORT NO. 36

JUNE 1986

PREPARED UNDER CONTRACT

N00014-77-C-0306 (NR-042-373)

FOR THE OFFICE OF NAVAL RESEARCH

Also prepared under National Science Foundation Grant MCS80-24649 and issued as Technical Report #250, Department of Statistics, Stanford University.

DEPARTMENT OF STATISTICS

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

Conditional Boundary Crossing Probabilities, with Applications to Change-Point Problems

by

Barry James

IMPA

Kang Ling James

IMPA

David Siegmund

Stanford University

For normal random walks $S_1, S_2, \cdots$, formed from independent, identically distributed random variables $X_1, X_2, \cdots$, we determine the asymptotic behavior under regularity conditions of

$$P(S_n > mg(n/m) \quad \text{for some} \quad n < m \mid S_m = m\xi_0, \quad U_m = m\lambda_0), \quad \xi_0 < g(1),$$

where $U_m = X_1^2 + \cdots + X_m^2$. The result is applied to a normal change-point problem to approximate null distributions of test statistics and to obtain approximate confidence sets for the change-point.

AMS 1980 subject classifications. Primary 60F10, 60J15. Secondary 62F03.

Key words and phrases. Boundary crossing probabilities, change-point, normal random walk.

1. **Introduction**. A method of developing approximations for boundary crossing probabilities which has received some attention of late is that of writing the probability as an expectation of a conditional boundary crossing probability given an appropriate random variable, and then developing an approximation for the conditional probability. Such a method has been used with some degree of success, as measured by the accuracy of the approximations, by Siegmund (1982, 1985, 1986), Hu (1985), and James, James, and Siegmund (1985).

Let $X_1, X_2, \cdots$ be independent, identically distributed $N(\mu, \sigma^2)$ random variables, with $S_n = X_1 + \cdots + X_n$ and $U_n = X_1^2 + \cdots + X_n^2$. Given a function $g(t)$, $0 < t \leq 1$, and $m \geq 1$, let τ be the possibly defective stopping time

$$\tau = \tau_m = \inf\{n \geq 1 : S_n > mg(n/m)\}$$

Siegmund (1982) studied the asymptotic behavior of the conditional probabilities

$$P(\tau < m \mid S_m = m\xi_0), \quad \xi_0 < g(1),$$

and used the results to approximate the tail probability of the Smirnov statistic and the power function of repeated significance tests for a normal mean when σ^2 is known. In this paper, we extend Siegmund's method to study the asymptotic order of the conditional probabilities

$$P(\tau < m \mid S_m = m\xi_0, \quad U_m = m\lambda_0), \quad \xi_0 < g(1), \tag{1.1}$$

and apply the result to some change-point problems.

Our main result is stated and proved in the next section, after some preliminary lemmas. The proof uses a likelihood ratio argument, but we believe the result could also be obtained using the method of Woodroofe (1982, Chapter 8). On the other hand, the method of mixtures of likelihood ratios (cf. Lai and Siegmund, 1977) and the method of Siegmund (1985, Theorem 9.54; see also Hu, 1985), which seem particularly simple in certain related problems, appear to be difficult to adapt to the present situation.

Our motivation for studying the conditional probabilities (1.1) comes from our investigation of the following change-point problem: Let $X_1, \cdots, X_m$ be independent random variables with $X_i \sim N(\mu_i, \sigma^2)$, and suppose we wish to test the hypothesis of no change in mean, H_0 :

$\mu_1 = \cdots = \mu_m$, versus the alternative of a single change, $H_1 : \mu_1 = \cdots = \mu_j \neq \mu_{j+1} = \cdots = \mu_m$ for some $j \in \{1, \cdots, m-1\}$. We can then use the theorem of the next section to obtain approximations for the significance levels of several tests of H_0, as well as to obtain likelihood-based confidence sets for the change-point j. These applications are given in Section 3.

2. **Asymptotic conditional boundary crossing probabilities.** Throughout this section, the following assumptions and definitions will hold. $X_1, \cdots, X_n$ are independent, identically distributed normal random variables, without loss of generality assumed to be $N(0,1)$, with $S_n = X_1 + \cdots + X_n$ and $U_n = X_1^2 + \cdots + X_n^2$, $n = 1, 2, \cdots, m$. The real-valued function g, defined on $(0,1]$, has two continuous derivatives. For a fixed $\xi_0 < g(1)$, there exists a unique point $t^* \in (0,1)$ which minimizes the function

$$h(t) = \frac{g(t) - \xi_0 t}{\{t(1-t)\}^{1/2}}$$

and further satisfies $h(t^*) > 0$, $\liminf_{t \to 0} h(t) > h(t^*)$, and $h''(t^*) > 0$. The stopping time $\tau = \tau_m$ is defined by

$$\tau = \inf \{n \leq m : S_n \geq mg(n/m)\};$$

we let $\tau = +\infty$ if the defining set is empty. Let λ_0 be such that $\lambda_0 > g^2(t^*)(t^*)^{-1} + \{g(t^*) - \xi_0\}^2(1-t^*)^{-1}$, and define μ and σ^2 by $\mu = g(t^*)/t^*$ and $\sigma^2 = \lambda_0 - g^2(t^*)(t^*)^{-1} - \{g(t^*) - \xi_0\}^2(1-t^*)^{-1}$. Let $\xi = m\xi_0$ and $\lambda = m\lambda_0$. Finally, for any $x \in \mathbb{R}$ and $y > 0$, we let

$$P_{x,y}^{(m)}(A) = P(A \mid S_m = x,\ U_m = y)$$

for A belonging to the σ-field generated by $X_1, \cdots, X_m$.

It can be seen that $\sigma^2 = \lambda_0 - \xi_0^2 - h^2(t^*)$, which in turn implies that $\lambda_0 > \xi_0^2$ and $\sigma^2/(\lambda_0 - \xi_0^2) < 1$. Note also that the condition $h'(t^*) = 0$ implies $\mu - g'(t^*) = (\mu - \xi_0)/\{2(1-t^*)\}$. Since $h(t^*) > 0$, this implies $\mu - g'(t^*) > 0$. It can also be shown that the conditions on h imply that $1 + 2g''(t^*)t^*(1-t^*)\{\mu - g'(t^*)\}^{-1} > 0$. Thus, the terms that appear in (2.6) in the statement of the theorem below are all well-defined, with the factor $\sigma^2/(\lambda_0 - \xi_0^2)$ taking on a value between 0 and 1.

The following two lemmas are technical and will be used in the proof of the theorem.

Lemma 1. Assume $a_m \to \infty$ with $a_m = o(m^{1/2})$, and let $b_m = m^{1/2}\log m$ and $I_m = (mt^* - a_m m^{1/2},\ mt^* + a_m m^{1/2})$. The following bounds all hold as $m \to \infty$:

(a) $\max_{1\le n\le m-1} P^{(m)}_{\xi,\lambda}(S_n \ge mg(n/m)) = O\left\{m^{-1/2}\left(\sigma^2/(\lambda_0-\xi_0^2)\right)^{(m-3)/2}\right\}$;

(b) $\sum_{n\in I_m} P^{(m)}_{\xi,\lambda}(S_n \ge mg(n/m)) = O\left\{\left(\sigma^2/(\lambda_0-\xi_0^2)\right)^{(m-3)/2}\right\}$;

(c) $P^{(m)}_{\xi,\lambda}\left(|\tau/m - t^*| \ge a_m m^{-1/2}\right) = o\left\{\left(\sigma^2/(\lambda_0-\xi_0^2)\right)^{(m-3)/2}\right\}$;

(d) $P^{(m)}_{\xi,\lambda}\left(|\tau/m - t^*| < a_m m^{-1/2},\ S_\tau - mg(\tau/m) \ge b_m\right) = o\left\{\left(\sigma^2/(\lambda_0-\xi_0^2)\right)^{(m-3)/2}\right\}$; and

(e) for each fixed $\epsilon > 0$, uniformly for n and r such that $|n - mt^*| \le a_m m^{1/2}$ and $0 \le r \le b_m$,

$$P^{(m)}_{\xi,\lambda}\left(|U_n/m - (\sigma^2+\mu^2)t^*| > \epsilon \mid S_n = mg(n/m) + r\right) = o(1).$$

Proof. (a) The conditional density of S_n given $S_m = \xi$ and $U_m = \lambda$, which is easily obtained via the conditional joint density of S_n and U_n, is given by

$$f_{S_n}(x \mid S_m = \xi,\ U_m = \lambda) = \left(\frac{m}{\pi n(m-n)}\right)^{1/2}\frac{\Gamma((m-1)/2)}{\Gamma((m-2)/2)} \cdot \left(\lambda - \frac{\xi^2}{m}\right)^{-(m-3)/2}\left(\lambda - \frac{(\xi-x)^2}{m-n} - \frac{x^2}{n}\right)^{(m-4)/2}$$

if $x^2n^{-1} + (\xi-x)^2(m-n)^{-1} < \lambda$ (and $= 0$ otherwise). After integrating and changing variables, we obtain

$$P^{(m)}_{\xi,\lambda}(S_n \ge mg(n/m)) = \frac{\Gamma((m-1)/2)}{\pi^{1/2}\Gamma((m-2)/2)}\int_{B_n}(1-y^2)^{(m-4)/2}dy,$$

where $B_n = \{y : |y| \le 1,\ y \ge (\lambda_0 - \xi_0^2)^{-1/2}h(n/m)\}$. Now if $0 < a < 1$ we can show, by a change of variables ($x = y^2$) and appropriate bounding of the integrand, that

$$\int_a^1 (1-y^2)^{(m-4)/2}dy \le \frac{(1-a^2)^{(m-2)/2}}{a(m-2)}. \tag{2.2}$$

Stirling's formula for the gamma function implies $\Gamma((m-1)/2)/\Gamma((m-2)/2) \sim (m/2)^{1/2}$. Thus, it follows from the fact that $h(n/m) \ge h(t^*)$, together with (2.1) and (2.2), that

$$P^{(m)}_{\xi,\lambda}(S_n \ge mg(n/m)) \le Km^{-1/2}\left(\frac{\lambda_0 - \xi_0^2 - h^2(n/m)}{\lambda_0 - \xi_0^2}\right)^{(m-3)/2} \tag{2.3}$$

for some $K > 0$ and all $m \ge 3$ and n such that $h^2(n/m) < \lambda_0 - \xi_0^2$ (the bound is 0 otherwise). Part (a) now follows from the relations $\sigma^2 = \lambda_0 - \xi_0^2 - h^2(t^*)$ and $h(n/m) \ge h(t^*)$.

(b) and (c). Note that

$$P_{\xi,\lambda}^{(m)}\left(|\tau/m - t^*| \geq a_m m^{-1/2}\right) \leq \sum_{n \notin I_m} P_{\xi,\lambda}^{(m)}\left(S_n \geq mg\left(n/m\right)\right),$$

and that (2.3) implies

$$(2.4) \quad P_{\xi,\lambda}^{(m)}\left(S_n \geq mg\left(n/m\right)\right) \leq Km^{-1/2}\left(\frac{\sigma^2}{\lambda_0 - \xi_0^2}\right)^{(m-3)/2}\left(\frac{\lambda_0 - \xi_0^2 - h^2(n/m)}{\lambda_0 - \xi_0^2 - h^2(t^*)}\right)^{(m-3)/2}.$$

Both (b) and (c) will follow by developing bounds for appropriate sums of the last factor above. By the assumptions on h, this factor will be of exponentially small order in m if n/m lies outside any fixed neighborhood of t^*; thus, for any fixed $\delta > 0$, we may restrict attention to n such that $|n/m - t^*| < \delta$. But Taylor's series expansions on $\log(1+x)$, to one derivative, and $h(t)$ around t^*, to two derivatives, yield the existence of $K_0 > 0$ and $\delta > 0$ such that

$$\left(\frac{\lambda_0 - \xi_0^2 - h^2(n/m)}{\lambda_0 - \xi_0^2 - h^2(t^*)}\right)^{(m-3)/2} < \exp\left(-mK_0\left(n/m - t^*\right)^2\right)$$

for n such that $|n/m - t^*| < \delta$. Parts (b) and (c) follow by summing these bounds over n in I_m and I_m^c and bounding the sums appropriately by integrals.

(d) By a process similar to that used to obtain (2.4), we have that

$$P_{\xi,\lambda}^{(m)}\left(|\tau/m - t^*| \leq a_m m^{-1/2},\ S_\tau - mg\left(\tau/m\right) \geq b_m\right) \leq \sum_{n \in I_m} P_{\xi,\lambda}^{(m)}\left(S_n \geq mg\left(n/m\right) + b_m\right)$$

and for some $K' > 0$, all $m \geq 3$, and all n such that $[h(n/m) + b_m/\{n(m-n)\}^{1/2}]^2 < \lambda_0 - \xi_0^2$,

$$\begin{aligned}
P_{\xi,\lambda}^{(m)}\left(S_n \geq mg\left(n/m\right) + b_m\right) &\leq Km^{-1/2}\left(\frac{\lambda_0 - \xi_0^2 - [h(n/m) + b_m/\{n(m-n)\}^{1/2}]^2}{\lambda_0 - \xi_0^2}\right)^{(m-3)/2} \\
&\leq K'm^{-1/2}\left(\frac{\sigma^2}{\lambda_0 - \xi_0^2}\right)^{(m-3)/2}\left(\frac{\lambda_0 - \xi_0^2 - h^2(n/m) - b_m^2/\{n(m-n)\}}{\lambda_0 - \xi_0^2 - h^2(t^*)}\right)^{(m-3)/2} \\
&\leq K'm^{-1/2}\left(\frac{\sigma^2}{\lambda_0 - \xi_0^2}\right)^{(m-3)/2}\left(1 - \frac{4b_m^2}{\sigma^2 m^2}\right)^{(m-3)/2},
\end{aligned}$$

where the last inequality uses the fact that $h(n/m) \geq h(t^*)$. Part (d) now follows by using the relation $1 - a \leq e^{-a}$ for $0 \leq a \leq 1$.

(e) By Markov's inequality

(2.5)

$$P^{(m)}_{\xi,\lambda}\left(|U_n/m-(\sigma^2+\mu^2)t^2|>\epsilon \mid S_n=x\right)\le \frac{1}{\epsilon^2}\Big[\mathrm{Var}\left(U_n/m \mid S_n=x,\ S_m=\xi,\ U_m=\lambda\right)$$
$$+\left\{E\left(U_n/m \mid S_n=x,\ S_m=\xi,\ U_m=\lambda\right)-(\sigma^2+\mu^2)t^*\right\}^2\Big].$$

Conditionally, U_n is a linear function of a beta-distributed random variable. In fact it can be shown that the random variable

$$V=\frac{U_n-S_n^2/n}{U_m-S_n^2/n-(S_m-S_n)^2/(m-n)},$$

whose numerator is one of two independent chi-squareds making up the denominator, has a beta distribution with parameters $(n-1)/2$ and $(m-n-1)/2$ and is independent of the vector (S_n,S_m,U_m). Therefore,

$$E(U_n \mid S_n=x,\ S_m=\xi,\ U_m=\lambda)=\frac{x^2}{n}+\left(\lambda-\frac{(\xi-x)^2}{m-n}-\frac{x^2}{n}\right)\frac{(n-1)}{(m-2)}$$

and

$$\mathrm{Var}(U_n \mid S_n=x,\ S_m=\xi,\ U_m=\lambda)=2\left(\lambda-\frac{(\xi-x)^2}{m-n}-\frac{x^2}{n}\right)^2\frac{(n-1)(m-n-1)}{m(m-2)^2}.$$

Part (e) now follows from (2.5) and the above by algebra. ∎

Lemma 2. For each $\epsilon>0$,

(a) $P^{(m)}_{\xi,\lambda}\left(|S_\tau/m-\mu t^*|>\epsilon,\ \tau<m\right)=o\left\{\left(\sigma^2/(\lambda_0-\xi_0^2)\right)^{(m-3)/2}\right\}$ and

(b) $P^{(m)}_{\xi,\lambda}\left(|U_\tau/m-(\sigma^2+\mu^2)t^*|>\epsilon,\ \tau<m\right)=o\left\{\left(\sigma^2/(\lambda_0-\xi_0^2)\right)^{(m-3)/2}\right\}$.

Proof. Let $a_m=\log m$. Applying first Lemma 1(c) and then the triangle inequality, we have

$$\begin{aligned}\mathrm{LHS}(a)&=P^{(m)}_{\xi,\lambda}\left(|S_\tau/m-\mu t^*|>\epsilon,\ |\tau/m-t^*|<a_m m^{-1/2}\right)+o\left\{\left(\frac{\sigma^2}{\lambda_0-\xi_0^2}\right)^{(m-3)/2}\right\}\\ &\le P^{(m)}_{\xi,\lambda}\left(S_\tau-mg(\tau/m)>\frac{m\epsilon}{2},\ |\tau/m-t^*|<a_m m^{-1/2}\right)\\ &\quad+P^{(m)}_{\xi,\lambda}\left(|g(\tau/m)-\mu t^*|>\frac{\epsilon}{2},\ |\tau/m-t^*|<a_m m^{-1/2}\right)+o\left\{\left(\frac{\sigma^2}{\lambda_0-\xi_0^2}\right)^{(m-3)/2}\right\}.\end{aligned}$$

The second summand on the right-hand side above is null for n sufficiently large, since g is continuous and $g(t^*)=\mu t^*$. The first summand can be handled by Lemma 1(d), thus completing the proof of (a).

Apply parts (c) and (d) of Lemma 1 to get

$$\mathrm{LHS}(b) = P_{\xi,\lambda}^{(m)}\left(|U_\tau/m - (\sigma^2+\mu^2)t^*| > \epsilon,\ S_\tau - mg\,(\tau/m) \le b_m,\ |\tau/m - t^*| < a_m m^{-1/2}\right) + o\left\{\left(\frac{\sigma^2}{\lambda_0 - \xi_0^2}\right)^{(m-3)/2}\right\}.$$

Now decompose the above event according to the value of τ, letting I_m be the interval $(mt^* - a_m\, m^{1/2},\ mt^* + a_m\, m^{1/2})$:

$$\begin{aligned}
\mathrm{LHS}(b) &\le \sum_{n\in I_m} P_{\xi,\lambda}^{(m)}\left(|U_n/m - (\sigma^2+\mu^2)t^*| > \epsilon,\ 0 \le S_n - mg\,(n/m) \le b_m\right) \\
&\quad + o\left\{\left(\frac{\sigma^2}{\lambda_0 - \xi_0^2}\right)^{(m-3)/2}\right\} \\
&\le \sum_{n\in I_m} P_{\xi,\lambda}^{(m)}\left(S_n \ge mg\,(n/m)\right) P_{\xi,\lambda}^{(m)}\left(|U_n/m - (\sigma^2+\mu^2)t^*| > \epsilon \mid 0 \le S_n - mg\,(n/m) \le b_m\right) \\
&\quad + o\left\{\left(\frac{\sigma^2}{\lambda_0 - \xi_0^2}\right)^{(m-3)/2}\right\}.
\end{aligned}$$

Part (b) now follows by applying first Lemma 1(e) and then Lemma 1(b). ∎

Remark 1. We will show in the course of the proof of our main theorem that $P_{\xi,\lambda}^{(m)}(\tau < m) > K\{\sigma^2/(\lambda_0 - \xi_0^2)\}^{(m-3)/2}$ for some $K > 0$. Then Lemmas 1(c) and 2 will give us convergence of τ/m, S_τ/m, and U_τ/m to t^*, μt^*, and $(\sigma^2+\mu^2)t^*$ in conditional $P_{\xi,\lambda}^{(m)}$-probability given $\{\tau < m\}$, i.e. for each $\epsilon > 0$

$$P_{\xi,\lambda}^{(m)}\left(|\tau/m - t^*| > \epsilon \mid \tau < m\right) \to 0,$$

$$P_{\xi,\lambda}^{(m)}\left(|S_\tau/m - \mu t^*| > \epsilon \mid \tau < m\right) \to 0,$$

and

$$P_{\xi,\lambda}^{(m)}\left(|U_\tau/m - (\sigma^2+\mu^2)t^*| > \epsilon \mid \tau < m\right) \to 0.$$

Remark 2. Formula (2.1) shows that the marginal probability $P_{\xi,\lambda}^{(m)}(S_n \ge mg(n/m))$ is maximized by that n which minimizes $h(n/m)$, i.e. by some n not far from mt^*. When m is large, then, it would seem reasonable that if the partial sum process were to cross the curve at all, it would do it for n near mt^*. We see from Remark 1 that this holds.

Theorem. Let ν be the function defined for $t > 0$ by

$$\nu(t) = 2t^{-2} \exp\left\{-2\sum_{n=1}^{\infty} n^{-1}\Phi(-tn^{1/2}/2)\right\},$$

where Φ is the standard normal distribution function. Then, as $m \to \infty$,

$$(2.6) \qquad P_{\xi,\lambda}^{(m)}(\tau < m) \sim \nu\left\{\frac{2(\mu - g'(t^*))}{\sigma}\right\}\left(\frac{\sigma^2}{\lambda_0 - \xi_0^2}\right)^{(m-3)/2}\left\{1 + \frac{2g''(t^*)t^*(1-t^*)}{\mu - g'(t^*)}\right\}^{-1/2}.$$

Remark 3. The function ν can be evaluated either directly by numerical computation or approximately, at least in the range $0 < t \leq 2$, from the local expansion

$$\nu(t) = \exp(-\rho t) + o(t^2), \qquad t \to 0,$$

where ρ is a numerical constant which is approximately equal to .583. See Siegmund (1985, Ch. X).

Proof. Let $P_{x,y,n}^{(m)}$ denote the restriction of $P_{x,y}^{(m)}$ to the σ-field generated by $X_1, \cdots, X_n$. Let $\xi_1 = m\mu$ and $\lambda_1 = m(\sigma^2 + \mu^2)$. The idea of the proof is to use a likelihood ratio argument, based on the likelihood ratio of $P_{\xi,\lambda}^{(m)}$ with respect to $P_{\xi_1,\lambda_1}^{(m)}$. The values of ξ_1 and λ_1 are chosen because of the approximately equivalent local behavior of the pre-τ process under $P_{\xi_1,\lambda_1}^{(m)}$ and, conditionally, under $P_{\xi,\lambda}^{(m)}$ given $\{\tau < m\}$. In fact, given $\{\tau < m\}$, Remark 1 tells us that $S_\tau/\tau \to \mu$ and $U_\tau/\tau \to \sigma^2 + \mu^2$ in $P_{\xi,\lambda}^{(m)}$-probability.

Let L_n denote the likelihood ratio of the absolutely continuous part of $P_{\xi,\lambda,n}^{(m)}$ relative to $P_{\xi_1,\lambda_1,n}^{(m)}$. A straightforward calculation shows that for $n \leq m - 2$,

$$L_n = \left(\frac{\lambda - U_n - (\xi - S_n)^2/(m-n)}{\lambda_1 - U_n - (\xi_1 - S_n)^2/(m-n)}\right)^{(m-n-3)/2}\left(\frac{\lambda_1 - \xi_1^2/m}{\lambda - \xi^2/m}\right)^{(m-3)/2}$$

if $\lambda_1 - U_n - (\xi_1 - S_n)^2/(m-n) > 0$ and $\lambda - U_n - (\xi - S_n)^2/(m-n) > 0$, and $L_n = 0$ if $\lambda - U_n - (\xi - S_n)^2/(m-n) \leq 0 < \lambda_1 - U_n - (\xi_1 - S_n)^2/(m-n)$.

By a slight generalization of Wald's likelihood ratio identity (see e.g. Siegmund 1985, p. 13),

$$(2.7) \qquad P_{\xi,\lambda}^{(m)}(\tau \leq m-2) = \int_{\{\tau \leq m-2\} \cap A} L_\tau \, dP_{\xi_1,\lambda_1}^{(m)} + P_{\xi,\lambda}^{(m)}(\{\tau \leq m-2\} \cap A^c),$$

where $A = \{\lambda_1 - U_\tau - (\xi_1 - S_\tau)^2/(m-\tau) > 0\}$. By Lemma 1(a), it is sufficient to show that the integral in (2.7) is asymptotically equivalent to the right-hand side of (2.6) and the final probability in (2.7) is of smaller asymptotic order.

Upon substitution of the likelihood ratio, the integral in (2.7) becomes

$$(2.8) \qquad \left(\frac{\sigma^2}{\lambda_0 - \xi_0^2}\right)^{(m-3)/2} \int_{\{\tau \le m-2\}\cap A} \left(\frac{\lambda - U_\tau - (\xi - S_\tau)^2/(m-\tau)}{\lambda_1 - U_\tau - (\xi_1 - S_\tau)^2/(m-\tau)}\right)^{(m-\tau-3)/2} dP^{(m)}_{\xi_1,\lambda_1}.$$

Law of large numbers arguments indicate that under $P^{(m)}_{\xi_1,\lambda_1}$, as $m \to \infty$,

$$\frac{\tau}{m} \xrightarrow{P} t^*, \quad \frac{S_\tau}{m} \xrightarrow{P} \mu t^*, \qquad \text{and} \quad \frac{U_\tau}{m} \xrightarrow{P} (\sigma^2 + \mu^2)t^*,$$

so that

$$(2.9)\ m^{-1}\left(\lambda - U_\tau - \frac{(\xi - S_\tau)^2}{m-\tau}\right) \xrightarrow{P} \sigma^2(1-t^*), \quad m^{-1}\left(\lambda_1 - U_\tau - \frac{(\xi_1 - S_\tau)^2}{m-\tau}\right) \xrightarrow{P} \sigma^2(1-t^*).$$

Since $\log(1+x) = x + O(x^2)$ as $x \to 0$, we have

(2.10)
$$\frac{\lambda - U_\tau - (\xi - S_\tau)^2/(m-\tau)}{\lambda_1 - U_\tau - (\xi_1 - S_\tau)^2/(m-\tau)} = \exp\Big[\frac{\lambda - \lambda_1 - \{\xi^2 - \xi_1^2 - 2S_\tau(\xi - \xi_1)\}/(m-\tau)}{\lambda_1 - U_\tau - (\xi_1 - S_\tau)^2/(m-\tau)} + O_p\left\{\left(\lambda_0 - \sigma^2 - \mu^2 - \frac{(\xi_0^2 - \mu^2 - 2(\xi_0 - \mu)S_\tau/m)}{1-\tau/m}\right)^2\right\}\Big].$$

Letting R_m be the excess over the boundary, i.e. $R_m = S_\tau - mg(\tau/m)$, and using a Taylor expansion on g at t^*, we get

$$(2.11)\ S_\tau = R_m + m\left\{g(t^*) + g'(t^*)(\tau/m - t^*) + \frac{g''(t^*)}{2}(\tau/m - t^*)^2 + \epsilon(\tau/m)(\tau/m - t^*)^2\right\},$$

where $\epsilon(t) \to 0$ as $t \to t^*$. To obtain the limiting joint distribution of R_m and $m^{1/2}(\tau/m - t^*)$ we must appeal to an appropriate nonlinear renewal theorem for the conditional process governed by $P^{(m)}_{\xi_1,\lambda_1}$. For an intuitive discussion of nonliner renewal theory which leads one to the correct limiting joint distribution, see Siegmund (1986, Appendix 2 and Lemma 2.16). Hu (1985, Chapter 4, Theorem 10) has proved a general result which provides a rigorous justification. The upshot is that R_m and $m^{1/2}(\tau/m - t^*)$ converge in distribution and are asymptotically independent under $P^{(m)}_{\xi_1,\lambda_1}$; the limiting distributions will be seen below. This, together with

some algebra, means that the right hand side of (2.10) can be written as

$$\exp\left\{\frac{2(\xi_0-\mu)(R_m+mg''(t^*)(\tau/m-t^*)^2/2+m\epsilon(\tau/m)(\tau/m-t^*)^2)}{(1-\tau/m)(\lambda_1-U_\tau-(\xi_1-S_\tau)^2/(m-\tau))}+O_p\left(m^{-2}\right)\right\}.$$

If we insert this in the integrand, (2.8) becomes

$$\left(\frac{\sigma^2}{\lambda_0-\xi_0^2}\right)^{(m-3)/2}\int_{\{\tau\le m-2\}\cap A}$$
$$\exp\left\{\frac{(m-\tau-3)(\xi_0-\mu)(R_m+mg''(t^*)(\tau/m-t^*)^2/2+o_p(1))}{(1-\tau/m)(\lambda_1-U_\tau-(\xi_1-S_\tau)^2/(m-\tau))}+O_p\left(m^{-1}\right)\right\}dP^{(m)}_{\xi_1,\lambda_1}.$$

Application of (2.9) then yields

$$\begin{aligned}(2.8)=&\left(\frac{\sigma^2}{\lambda_0-\xi_0^2}\right)^{(m-3)/2}\int_{\{\tau\le m-2\}\cap A}\\ &\exp\left\{\frac{(\xi_0-\mu)(R_m+mg''(t^*)(\tau/m-t^*)^2/2)}{\sigma^2(1-t^*)}+o_p(1)\right\}dP^{(m)}_{\xi_1,\lambda_1}.\end{aligned}\tag{2.12}$$

It follows from (2.9) that

$$P^{(m)}_{\xi_1,\lambda_1}\left(\tau\le m-2,\ \lambda_1-U_\tau-\frac{(\xi_1-S_\tau)^2}{m-\tau}>0\right)\to 1,\qquad m\to\infty,$$

so that if we may interchange expectation and limit in (2.12), we will be able to evaluate the order of (2.8) by using Hu's result. This result states that as $m\to\infty$

$$P^{(m)}_{\xi_1,\lambda_1}\left(\frac{(\tau-mt^*)(\mu-g'(t^*))}{(mt^*(1-t^*))^{1/2}\sigma}\le x,\ R_m\le y\right)\to\Phi(x)\cdot\lim_{c\to\infty}P(R_c^*\le y),$$

where Φ is the standard normal distribution function and R_c^* is the excess over the constant boundary c of a random walk, generated by independent, identically distributed $N(\mu-g'(t^*),\ \sigma^2)$ random variables, which is stopped the first time it exceeds c. If R is a random variable such that $R_c^*\xrightarrow{\mathcal{D}}R$ as $c\to\infty$, then renewal theory (see Siegmund, 1985, Chapter VIII) allows us to calculate

$$E\exp\left\{\frac{(\xi_0-\mu)R}{(1-t^*)\sigma^2}\right\}=\nu\left\{\frac{2(\mu-g'(t^*))}{\sigma}\right\},$$

where we use the fact, noted earlier in this section, that $\mu-g'(t^*)=(\mu-\xi_0)/\{2(1-t^*)\}$. If X has a chi-squared distribution with one degree of freedom, the remaining factor will have the form

$$E\exp\left\{-\frac{g''(t^*)t^*(1-t^*)}{\mu-g'(t^*)}X\right\}=\left\{1+\frac{2g''(t^*)t^*(1-t^*)}{\mu-g'(t^*)}\right\}^{-1/2}.$$

Therefore, we will be able to conclude that the right-hand side of (2.6) is asymptoticaly equivalent to (2.8) if we can make the exchange of expectation and limit alluded to above.

Fatou's Lemma for convergence in law implies that the right-hand side of (2.6) is an asymptotic lower bound for (2.8). By Lemma 1(c), $P^{(m)}_{\xi,\lambda}(m^{1/2}|\tau/m - \tau^*| \geq a_m)$ is of asymptotically smaller order than this if $a_m \to \infty$ and $a_m = o(m^{1/2})$. The analog of formula (2.7) with $\{m^{1/2}|\tau/m - t^*| \geq a_m\}$ in place of $\{\tau \leq m-2\}$ then implies that

$$\int_{\{m^{1/2}|\tau/m-t^*|\geq a_m\}\cap A} L_\tau dP^{(m)}_{\xi_1,\lambda_1} = o\left\{\left(\frac{\sigma^2}{\lambda_0 - \xi_0^2}\right)^{(m-3)/2}\right\}.$$

Therefore, for any sequence $\{a_m\}$ such that $a_m \to \infty$ and $a_m = o(m^{1/2})$,

$$(2.13)\qquad \begin{aligned}(2.8) \sim \left(\frac{\sigma^2}{\lambda_0 - \xi_0^2}\right)^{(m-3)/2} &\int_{\{m^{1/2}|\tau/m-t^*|<a_m\}\cap A} \\ &\left(\frac{\lambda - U_\tau - (\xi - S_\tau)^2/(m-\tau)}{\lambda_1 - U_\tau - (\xi_1 - S_\tau)^2/(m-\tau)}\right)^{(m-\tau-3)/2} dP^{(m)}_{\xi_1,\lambda_1}.\end{aligned}$$

From Lemmas 1 and 2, via Remark 1, we see that for all $\epsilon > 0$

$$(2.14)\qquad P^{(m)}_{\xi,\lambda}\left(\left|m^{-1}\left\{\lambda_1 - U_\tau - \frac{(\xi_1 - S_\tau)^2}{m-\tau}\right\} - \sigma^2(1-t^*)\right| > \epsilon \mid \tau \leq m-2\right) \to 0,$$

so that we may replace the event A in the above by $\{(\lambda_1 - U_\tau - (\xi_1 - S_\tau)^2/(m-\tau))/m > \sigma^2(1-t^*)/2\}$. We may use the fact that $\log(1+x) < x$ if $x > 0$ to obtain bounds for the integrand in the right-hand side of (2.13) in the region of integration:

$$\begin{aligned}&\left(\frac{\lambda - U_\tau - (\xi - S_\tau)^2/(m-\tau)}{\lambda_1 - U_\tau - (\xi_1 - S_\tau)^2/(m-\tau)}\right)^{(m-\tau-3)/2} \\ &\leq 1 \vee \exp\left[\frac{(m-\tau-3)}{2}\left(\frac{\lambda - \lambda_1 - \{\xi^2 - \xi_1^2 - 2S_\tau(\xi - \xi_1)\}/(m-\tau)}{\lambda_1 - U_\tau - (\xi_1 - S_\tau)^2/(m-\tau)}\right)\right] \\ &\overset{\text{by}(2.11)}{\leq} 1 \vee \exp\left[\frac{2(m-\tau-3)(\xi_0 - \mu)\{R_m + mg''(t^*)(\tau/m - t^*)^2/2 + m\epsilon(\tau/m)(\tau/m) - t^*)^2\}}{\sigma^2(1-t^*)(1-\tau/m)}\right] \\ &\overset{m\text{ large}}{\leq} 1 \vee \exp\left[\frac{m(\tau/m - t^*)^2\{(\xi_0 - \mu)g''(t^*) + 2|\epsilon(\tau/m)|\}}{\sigma^2(1-t^*)}\right] \leq \exp(Ka_m^2),\end{aligned}$$

for some $K > 0$ and m large enough to make $a_m/m^{1/2}$ sufficiently small. We may then choose a suitable sequence $\{a_m\}$, converging sufficiently slowly to $+\infty$, to allow us to finish the proof of the asymptotic equivalence of (2.8) and the right-hand side of (2.6).

Equation (2.14) also indicates that the final probability in (2.7) is of smaller asymptotic order, because it implies that $P^{(m)}_{\xi,\lambda}(A^c \mid \tau \leq m-2) \to 0$. ∎

As an aid in applying the theorem, we note that if g has two continuous derivatives, then the other conditions on g and h will be satisfied if (i) t^* is the only point at which $h' = 0$, (ii) $h(t^*) > 0$, and (iii) $h''(t^*) > 0$. This will be the case in the examples considered in the next section.

3. Application: Tests and confidence sets for a change-point.

3.1. Approximate significance levels. Let X_i, $i = 1, \cdots, m$, be independent normal random variables with unknown means μ_i and constant unknown variance $\sigma^2 > 0$. We consider testing the hypothesis of a constant mean against the alternative of a single change-point, i.e. testing $H_0 : \mu_1 = \cdots = \mu_m$ versus H_1 : for some $j \in \{1, \cdots, m-1\}$, $\mu_1 = \cdots = \mu_j \neq \mu_{j+1} = \cdots = \mu_m$.

We will focus here on three tests considered in James, James, and Siegmund (1985): the *likelihood ratio test*, a *Studentized version of a score-like test due to* Pettitt (1980) which was introduced for the case of known variance, and a modification of the recursive residuals test proposed by Brown, Durbin and Evans (1975). Approximations to the significance levels of these tests were given, without formal justification, in James, James, and Siegmund (1985). We will now show how the theorem of the last section can be used to obtain these approximations and also obtain approximate confidence sets for the change-point j. We first give the test statistics for the three tests being considered.

The generalized likelihood ratio test can be easily shown to be based on the statistic

$$\max_{1 \leq n \leq m-1} \frac{|S_n - n\bar{X}_m|}{\{n(1-n/m)\}^{1/2} S},$$

where $S_n = X_1 + \cdots + X_n$, $\bar{X}_n = (X_1 + \cdots + X_n)/n$, and $S = \left\{m^{-1} \sum_{n=1}^{m} (X_n - \bar{X}_m)^2\right\}^{1/2}$. In the 1985 paper cited above, we considered the larger family of tests based on statistics of the form

$$T_1 = \max_{m_0 \leq n \leq m_1} \frac{|S_n - n\bar{X}_m|}{\{n(1-n/m)\}^{1/2} S},$$

where $1 \leq m_0 < m_1 \leq m-1$. The use of T_1 with $m_0 > 1$ and $m_1 < m-1$ allows one greater power than that of the likelihood ratio test for values of j near $m/2$, while giving up some

power for small and large j where the change is difficult to detect in any case. If $m_0 = m - m_1$, then the test based on T_1 belongs to a general family of tests considered by Deshayes and Picard (1984).

The Studentized version of Pettitt's test is based on the statistic

$$T_2 = \max_{1 \le n \le m-1} \frac{|S_n - n\bar{X}_m|}{S}.$$

The recursive residuals statistic of Brown, Durbin and Evans is formed, in the case of known variance, by accumulating sums of standardized residuals Z_n of the X_{n+1} from the previous means $\bar{X}_n$, that is

$$Z_n = \{n/(n+1)\}^{1/2}(X_{n+1} - \bar{X}_n), \quad n = 1, 2, \cdots, m-1.$$

The Z_n are independent with common distribution $N(0, \sigma^2)$ under H_0. In the case of unknown variance, the accumulated sums are Studentized by dividing by the sample standard deviation of the Z's. In James, James, and Siegmund (1985), power considerations lead us to study the statistic obtained by summing the standardized residuals "from the right," so that we base the recursive residuals test on the statistic

$$T_3 = \max_{m_0 \le n \le m-1} \frac{|S'_n|}{S' n^{1/2}},$$

where $S'_n = Z_{m-1} + \cdots + Z_{m-n}$ and $S' = \{(m-1)^{-1}(Z_1^2 + \cdots + Z_{m-1}^2)\}^{1/2}$. However, as long as we are only concerned with the significance level, it makes no difference whether the recursive residuals are summed from the right or, as Brown, Durbin, and Evans proposed, from the left. We note also that although we are considering this very simple problem of a change in mean, the null hypothesis distribution of the recursive residuals statistic is in fact the same as in the general regression model of Brown, Durbin, and Evans, the only difference being that there are $m - p$ recursive residuals Z_n instead of $m - 1$, where p is the number of regression parameters.

We will first obtain approximate significance levels for the tests based on T_1 and T_2. Since the distribution of the process $\{(S_n - n\bar{X}_m)/S,\ n = 0, 1, \cdots, m\}$ does not depend on (μ, σ^2), the process is independent of the complete sufficient statistic (S_m, U_m), by Basu's theorem

(Lehmann, 1959, Theorem 5.2). Therefore, starting with the easier to handle T_2 we have

$$P(T_2 \geq b) = P_{0,m}^{(m)}(T_2 \geq b) = P_{0,m}^{(m)}(\max_{1 \leq n \leq m-1} |S_n| \geq b).$$

If we assume that $b = b_m = m\varsigma$, for some $0 < \varsigma < \frac{1}{2}$, then we can apply our theorem with $g(t) \equiv \varsigma$, $\xi_0 = 0$, $\lambda_0 = 1$, and $t^* = \frac{1}{2}$ to obtain

$$P(T_2 \geq b) \sim 2\nu \left\{ \frac{4\varsigma}{(1 - 4\varsigma^2)^{1/2}} \right\} (1 - 4\varsigma^2)^{(m-3)/2}. \tag{3.1}$$

An approximate size α test based on T_2 can now be obtained by using as critical value $b = m\varsigma$, where ς makes the right-hand side of (3.1) equal to α.

For the modified likelihood ratio test, we have

$$\begin{aligned} P(T_1 \geq b) &= P_{0,m}^{(m)}(T_1 \geq b) \\ &= P_{0,m}^{(m)} \left(\max_{m_0 \leq n \leq m_1} \frac{|S_n|}{\{n(1 - n/m)\}^{1/2}} \geq b \right). \end{aligned}$$

Conditioning with respect to the values of S_{m_1} and U_{m_1} and using the Markov property, we have

$$\begin{aligned} P(T_1 \geq b) &= P_{0,m}^{(m)} \left(|S_{m_1}| \geq b \{m_1 (1 - m_1/m)\}^{1/2} \right) \\ &\quad + \int \int_{A_m} P_{x,y}^{(m_1)}(\tau < m_1) P_{0,m}^{(m)}(S_{m_1} \in dx,\, U_{m_1} \in dy), \end{aligned} \tag{3.2}$$

where $\tau = \inf[n \geq m_0 : S_n \geq b\{n(1 - n/m)\}^{1/2}]$ and A_m is the set of (x, y) such that $|x| < b\{m_1(1 - m_1/m)\}^{1/2}$ and the $P_{0,m}^{(m)}$-joint density of S_{m_1} and U_{m_1}, as a function of x and y, is positive.

The first summand in the right-hand side of (3.2) can be calculated exactly, as in (2.1). The theorem can be used to approximate the integrand in the second summand. For this, assume that $b = cm^{1/2}$, $m_0 = mt_0$, $m_1 = mt_1$, $x = m_1 x_0$, and $y = m_1 y_0$, where $0 < c < 1$, $0 < t_0 < t_1 < 1$, $|x_0| < ct_1^{-1/2}(1 - t_1)^{1/2}$, and $m_1 y_0$ is a $P_{0,m}^{(m)}$-possible value of U_{m_1}, given $S_{m_1} = m_1 x_0$. For $g(t) = c\{tt_1^{-1}(1 - tt_1)\}^{1/2}$, $\xi_0 = x_0$ and $\lambda_0 = y_0$, we obtain

$$t^* = \frac{x_0^2 t_1}{c^2(1 - t_1)^2 + x_0^2 t_1^2}.$$

The only hitch in applying the theorem is that $\tau \geq m_0$; a direct application requires $m_0 = 1$. However, this is no problem if $t^* > t_0/t_1$, because in this case it follows from (2.4) that the

$P_{x,y}^{(m_1)}$-probability of the process's crossing the boundary before $n = m_0$ is of exponentially smaller order than that of crossing before $n - m_1$, so that we may replace m_0 by 1. On the other hand, if $t^* < t_0/t_1$, which corresponds to $|x_0| < ct_1^{-1}(1-t_1)\{t_0(1-t_0)^{-1}\}^{1/2}$, we can approximate the integrand by 0. This follows again from (2.4), which implies that the integrand will be of exponentially smaller order than other values of the integrand corresponding to $t^* > t_0/t_1$. Therefore, we are led to the approximation

$$
\begin{aligned}
P(T_1 \geq b) \cong & \frac{2\Gamma((m-1)/2)}{\pi^{1/2}\Gamma((m-2)/2)} \int_c^1 (1-x^2)^{(m-4)/2} dx \\
& + 2c\left(\frac{1-t_1}{t_1}\right)^{1/2} \int\int_B \frac{1}{x_0} \nu\left(\frac{x_0}{t^*(1-t_1)\sigma}\right)\left(\frac{\sigma^2}{y_0 - x_0^2}\right)^{(m_1-3)/2} \\
& \cdot P_{0,m}^{(m)}\left(\frac{S_{m_1}}{m_1} \in dx_0,\ \frac{U_{m_1}}{m_1} \in dy_0\right),
\end{aligned}
$$

where $\sigma^2 = \sigma^2(x_0, y_0) = y_0 - c^2 t_1^{-1} + x_0^2 t_1(1-t_1)^{-1}$ and

$$
\begin{aligned}
B = \Big\{(x_0, y_0) : & \frac{c^2}{t_1} - \frac{x_0^2 t_1}{1-t_1} \leq y_0 \leq \frac{1}{t_1} - \frac{x_0^2 t_1}{1-t_1}, \frac{c(1-t_1)}{t_1}\left(\frac{t_0}{1-t_0}\right)^{1/2} \\
& < x_0 < c\left(\frac{1-t_1}{t_1}\right)^{1/2}\Big\}.
\end{aligned}
$$

The factor 2 above is due to restriction to positive values of x, by symmetry.

A further approximaton can be made upon insertion of the conditional density into the integral, with subsequent utilization of the fact that U_{m_1} is conditionally, given $S_m = 0$, $U_m = m$, and $S_{m_1} = m_1 x_0$, a linear function of a random variable with a beta distribution with parameters $(m_1 - 1)/2$ and $(m - m_1 - 1)/2$, which as $m \to \infty$ with $m_1/m \to t_1$ collapses to a point mass at t_1. Following this procedure, we can insert the $P_{0,m}^{(m)}$-density of $(S_{m_1}/m_1,\ U_{m_1}/m_1)$, to wit

$$
\frac{\{t_1^{m_1}/(1-t_1)\}^{1/2}\Gamma((m-1)/2)}{\pi^{1/2}\Gamma((m_1-1)/2)\Gamma((m-m_1-1)/2)}(y_0 - x_0^2)^{(m_1-3)/2}[1 - t_1\{y_0 + x_0^2 t_1/(1-t_1)\}]^{(m-m_1-3)/2},
$$

make the change of variable (in y_0)

$$
z = t_1(1-c^2)^{-1}\{y_0 + x_0^2 t_1(1-t_1)^{-1} - c^2 t_1^{-1}\},
$$

integrate with respect to z, and use Stirling's formula to approximate the remaining gamma functions, to show that the double integral appearing above reduces asymptotically to a single

integral in x_0. Thus we are led to the approximation

$$(3.3)\qquad \begin{aligned} P(T_1 \geq b) \cong \left(\frac{2m}{\pi}\right)^{1/2} \int_c^1 (1-x^2)^{(m-4)/2} dx \\ + c\left(\frac{2m}{\pi}\right)^{1/2} (1-c^2)^{(m-4)/2} \int_{c\{(t_1^{-1}-1)/(1-c^2)\}^{1/2}}^{c\{(t_0^{-1}-1)/(1-c^2)\}^{1/2}} \frac{1}{x}\nu\left\{x + \frac{c^2}{(1-c^2)x}\right\} dx. \end{aligned}$$

<u>Remark 4</u>. It is easy to see that for each $i = 0, 1, \cdots$

$$\int_c^1 (1-x^2)^{(m-i)/2} dx = (cm)^{-1}(1-c^2)^{(m-i+2)/2}\{1 + m^{-1}(i - 1 - c^{-2}) + o(m^{-1})\}$$

as $m \to \infty$. Use of this approximation simplifies slightly the computational burden associated with application of (3.3) or (3.4) below. From this expansion it is evident that the first term on the right-hand side of (3.3) is asymptotically of smaller order than the second and mathematically could be neglected. In a number of related problems Siegmund (1985) shows numerically that including this term typically improves the approximation, and hence we have included it for numerical purposes.

<u>Remark 5</u>. It is natural to ask what precise mathematical meaning can be attached to (3.3). As noted in Remark 4, the first integral on the right hand side of (3.3) is asymptotically negligible. With some additional work it can be shown that $P(T_1 \geq b)$ is asymptotically equivalent to the second integral on the right-hand side of (3.3). It suffices to show that for each $x_0 \in (ct_1^{-1}(1-t_1)\{t_0/(1-t_0)\}^{1/2}, c\{(1-t_1)/t_1\}^{1/2})$ the asymptotic behavior of the conditional probability indicated in the Theorem holds uniformly for y_0 in a neighborhood of $1 + c^2 t_1^{-1}(1-t_1) - x_0^2 t_1 (1-t_1)^{-1}$ of width $a_m/m^{1/2}$, where $a_m \to \infty$, together with appropriate uniformity in Lemma 1. The details are tedious and have been omitted.

The procedure in studying T_3 is similar to that of T_1. Starting with sufficiency arguments, we have

$$\begin{aligned} P(T_3 \geq b) = P(T_3 \geq b \mid (S')^2 = 1) = P(|S'_{m-1}| \geq b(m-1)^{1/2} \mid (S')^2 = 1) \\ + \int_{-b(m-1)^{1/2}}^{b(m-1)^{1/2}} P(T_3 \geq b \mid S'_{m-1} = x,\ (S')^2 = 1) P(S'_{m-1} \in dx \mid (S')^2 = 1). \end{aligned}$$

The first summand on the right-hand side above can be calculated exactly from the conditional density. The last summand can be approximated by using the theorem to approximate the integrand. In this case, we assume $b = c(m-1)^{1/2}$, $x = (m-1)x_0$, and $m_0 = (m-1)t_0$, where

$0 < c < 1$, $|x_0| < c$, and $0 < t_0 < 1$, and apply the theorem with $g(t) = ct^{1/2}$, $\xi_0 = x_0$, $\lambda_0 = 1$, and $t^* = x_0^2/c^2$. Calculations similar to those done for the likelihood ratio test then lead to the approximation

$$
\begin{aligned}
P(T_3 \geq b) \cong &\left\{\frac{2(m-1)}{\pi}\right\}^{1/2} \int_c^1 (1-x^2)^{(m-4)/2} dx \\
&+ c\left(\frac{2(m-1)}{\pi}\right)^{1/2} (1-c^2)^{(m-4)/2} \int_{c/(1-c^2)^{1/2}}^{c/\{t_0(1-c^2)\}^{1/2}} x^{-1}\nu(x)dx.
\end{aligned}
\tag{3.4}
$$

Some comments on the numerical accuracy of the above approximations can be found in James, James, and Siegmund (1985).

3.2. **Confidence sets for the change-point.** Our theorem can be used to obtain approximate, likelihood-based confidence sets for j, assuming that it exists. The method extends that of Siegmund (1986, §3.5), who considered the case of known variance.

Suppose $X_1, \cdots, X_m$ are independent, with $X_1, \cdots, X_j$ independent, identically distributed $N(\mu_1, \sigma^2)$ and $X_{j+1}, \cdots, X_m$ independent, identically distributed $N(\mu_2, \sigma^2)$ for some $1 \leq j \leq m-1$, $\mu_1 \neq \mu_2$, and $\sigma^2 > 0$ unknown. To test $H_\rho : j = \rho$ versus $K_\rho : j \neq \rho$, the likelihood ratio test can be based on the statistic

$$T_\rho = \max_{k \neq \rho} \frac{W_k - W_\rho}{\sum_{i=1}^m (X_i - \bar{X}_m)^2 - W_\rho},$$

where $W_k = (S_k - k\bar{X}_m)^2/\{k(1-k/m)\}$. Under H_ρ, the distribution of T_ρ depends only on ρ and $\delta/\sigma = (\mu_2 - \mu_1)/\sigma$, and if we actually wanted to perform the test, or obtain a confidence set based on the family of such tests, we would need to find $c = c_\alpha$ such that $\alpha = \sup\{P_{\rho,\delta/\sigma}(T_\rho \geq c) : \delta/\sigma \neq 0\}$. This problem seems impracticable. However, since T_ρ is stochastically independent of S_m we do have

$$P_{\rho,\delta/\sigma}(T_\rho \geq c) = P_{\rho,\delta/\sigma}(T_\rho \geq c \mid S_m = 0),$$

which has the effect of eliminating $\bar{X}_m$ from the expression for T_ρ and putting it as a conditioner. In the spirit of Siegmund (1986, §3.5), we can then think of performing the test conditionally by conditioning first on S_ρ, which eliminates the dependence on δ/σ, and then on U_ρ and

$U_m - U_\rho$. That is, we attempt to find $c = c(\alpha, \rho, \xi, \lambda_1, \lambda_2)$ such that

$$\alpha = P_\rho(T_\rho \geq c \mid S_\rho = \xi, S_m = 0, U_\rho = \lambda_1, U_m - U_\rho = \lambda_2)$$

$$(3.5) \qquad = P_\rho\Big(|S_k| \geq \left\{c\left(\lambda_1 + \lambda_2 - \frac{\xi^2}{\rho(1-\rho/m)}\right) + \frac{\xi^2}{\rho(1-\rho/m)}\right\}^{1/2} \{k(1-k/m)\}^{1/2}$$
$$\text{for some} \quad k \neq \rho \mid S_\rho = \xi, S_m = 0, U_\rho = \lambda_1, U_m - U_\rho = \lambda_2\Big).$$

A likelihood based, level $1 - \alpha$ confidence set $C(X)$ could then be based on the family of size α tests determined by (3.5), by defining

$$C(X) = \left\{\rho : T_\rho < c\left(\alpha, \rho, S_\rho - \rho\bar{X}_m, \sum_{k=1}^{\rho}(X_k - \bar{X}_m)^2, \sum_{k=\rho+1}^{m}(X_k - \bar{X}_m)^2\right)\right\}.$$

The above procedure can be carried out by using the theorem to obtain approximate values of $c(\alpha, \rho, \xi, \lambda_1, \lambda_2)$. Assume $\rho/m \to p$, $\xi = m\xi_0$, $\lambda_1 = m\lambda_{10}$, and $\lambda_2 = m\lambda_{20}$, where $0 < p < 1$, ξ_0, λ_{10}, and λ_{20} are fixed. Let $c_0 = \{c(\lambda_{10} + \lambda_{20}) + (1-c)\xi_0^2 p^{-1}(1-p)^{-1}\}$. Then the Markov property implies

$$\text{RHS (3.5)} = p_1 + p_2 - p_1 p_2,$$

where

$$p_1 = P_\rho\left(|S_k| \geq mc_0\left\{\frac{k}{m}\left(1 - \frac{k}{m}\right)\right\}^{1/2} \quad \text{for some} \quad k < \rho \mid S_\rho = \xi, U_\rho = \lambda_1\right)$$

and

$$p_2 = P_{m-\rho}\left(|S_k| \geq mc_0\left\{\frac{k}{m}\left(1 - \frac{k}{m}\right)\right\}^{1/2} \quad \text{for some} \quad k < m - \rho \mid S_{m-\rho} = \xi, U_{m-\rho} = \lambda_2\right).$$

Applying the theorem with $g(t) = c_0\{t(1-pt)/p\}^{1/2}$, we obtain

$$p_1 \sim \nu\left(\frac{\xi_0}{(1-p)t_1^*\sigma_1}\right)\left(\frac{\sigma_1^2}{\lambda_{10} - \xi_0^2}\right)^{(\rho-3)/2} \frac{c_0}{\xi_0}\left(\frac{1-p}{p}\right)^{1/2}$$

and

$$p_2 \sim \nu\left(\frac{\xi_0}{pt_2^*\sigma_2}\right)\left(\frac{\sigma_2^2}{\lambda_{20} - \xi_0^2}\right)^{(m-\rho-3)/2} \frac{c_0}{\xi_0}\left(\frac{p}{1-p}\right)^{1/2},$$

where $t_1^* = \xi_0^2 p\{c_0^2(1-p)^2 + \xi_0^2 p^2\}^{-1}$, $t_2^* = \xi_0^2(1-p)\{c_0^2 p^2 + \xi_0^2(1-p^2)\}^{-1}$, $\sigma_1^2 = \lambda_{10} - c_0^2 p^{-1} - \xi_0^2 p(1-p)^{-1}$, and $\sigma_2^2 = \lambda_{20} - c_0^2(1-p)^{-1} - \xi_0^2(1-p)p^{-1}$ (σ_1^2 and σ_2^2 are assumed > 0).

References

Brown, R. L., Durbin, J. and Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time. *J. Roy. Statist. Soc.* B **37**, 149–192.

Deshayes, J. and Picard, D. (1984). Lois asymptotiques des tests et estimateurs de rupture dans un modèle statistique classique. *Ann. Inst. Henri Poincaré* **20**, 309–327.

Hu, I. (1985). Repeated significance tests for exponential families. Technical Report No. 34. Department of Statistics, Stanford University.

James, B., James, K. L. and Siegmund, D. (1985). Tests for a change-point. Technical Report No. 35. Department of Statistics, Stanford University.

Lai, T. L. and Siegmund, D. (1977). A nonlinear renewal theory with applications to sequential analysis I. *Ann. Statist.* **5**, 946–954.

Lehmann, E. L. (1959). *Testing Statistical Hypotheses.* John Wiley and Sons, New York.

Pettitt, A. N. (1980). A simple cumulative sum type statistic for the change-point problem with zero-one observations. *Biometrika* **67**, 79–84.

Siegmund, D. (1982). Large deviations for boundary crossing probabilities. *Ann. Probab.* **10**, 581–588.

Siegmund, D. (1985). *Sequential Analysis: Tests and Confidence Intervals.* Springer-Verlag, New York-Heidelberg-Berlin.

Siegmund, D. (1986). Boundary crossing probabilities and statistical applictions. *Ann. Statist.* **14**.

Woodroofe, M. (1982). *Nonlinear Renewal Theory in Sequential Analysis.* SIAM, Philadelphia.

END
DTIC
8-86